DR RUI XIA (Orcid ID : 0000-0003-2409-1181)

Article type : Research review

*Research review*

# MicroRNA482/2118, a miRNA superfamily essential for both disease resistance and plant development

Yanqing Zhang[1,2,3,4,†], Muhammad Waseem[1,2,3,4,†], Zaohai Zeng[1,2,3,4], Jing Xu[4], Chengjie Chen[1,2,3,4], Yuanlong Liu[1,2,3,4], Jixian Zhai[5], Rui Xia[1,2,3,4,*]

[1]State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources; South China Agricultural University, Guangzhou, Guangdong 510640, China;

[2]Guangdong Laboratory for Lingnan Modern Agriculture; South China Agricultural University, Guangzhou, Guangdong 510640, China;

[3]Key Laboratory of Biology and Germplasm Enhancement of Horticultural Crops in South China, Ministry of Agriculture and Rural Affairs, South China Agricultural University, Guangzhou, Guangdong 510640, China;

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/NPH.17853

[4]Guangdong Litchi Engineering Research Center, College of Horticulture, South China Agricultural University, Guangzhou, Guangdong 510640, China;

[5]Institute of Plant and Food Science, Department of Biology, Southern University of Science and Technology, Shenzhen 518055, China.

**ORCID IDs:**

Yanqing Zhang: 0000-0002-5343-7404

Muhammad Waseem: 0000-0001-7947-7722

Chengjie Chen: 0000-0001-5964-604X

Yuanlong Liu: 0000-0003-2907-9069

Jixian Zhai: 0000-0002-0217-0666

Rui Xia: 0000-0003-2409-1181

† These authors contribute equally to the review.

* Correspondence author: Dr. Rui Xia, rxia@scau.edu.cn

Received: *10 October 2021*

Accepted: *7 November 2021*

**Summary**

MicroRNAs (miRNAs) are a class of 21-24 nucleotides (nt) non-coding small RNAs ubiquitously distributed across the plant kingdom. miR482/2118, one of the conserved miRNA superfamilies originating from gymnosperms, has divergent main functions in core-angiosperms. It mainly regulates *NUCLEOTIDE BINDING SITE-LEUCINE-RICH REPEAT* (*NBS-LRR*) genes in eudicots, functioning as an essential component in plant disease resistance; in contrast, it predominantly targets numerous long non-coding RNAs (lncRNAs) in monocot grasses, which are vital for plant reproduction. Usually, miR482/2118 is 22-nt in length, which can trigger the production of phased small interfering RNAs (phasiRNAs) after directed cleavage. PhasiRNAs instigated from target genes of miR482/2118 enhance their roles in corresponding biological processes by *cis*-regulation on cognate genes and expands their function to other pathways via *trans* activity on different genes. This review summarizes the origin, biogenesis, conservation, and evolutionary characteristics of the miR482/2118 superfamily and delineates its diverse functions in disease resistance, plant development, stress responses, etc.

**Keywords:** miR482/2118 superfamily, phasiRNAs, evolution, function divergence, reproduction, disease resistance, NBS-LRR, genic male sterility

## Ⅰ. Introduction

MicroRNAs (miRNAs) are a large class of small non-coding RNAs of 21-24 nucleotide (nt) in length and a master regulator of indispensable biological processes (D'Ario *et al.*, 2017; Chen *et al.*, 2018). For miRNA biogenesis, a *MIRNA* gene is transcribed by RNA polymerase II (Pol II) into a primary miRNA transcript (pri-miRNA), which forms a stem-loop structure. It is then processed into a miRNA/miRNA* duplex by RNase III family enzyme DICER-LIKE1 (DCL1). The duplex is protected from degradation by 3' end methylation induced by HUA ENHANCER1 (HEN1) (Yu *et al.*, 2005). The mature miRNA strand is incorporated into ARGONAUTE 1 protein (AGO1) in the RNA-induced silencing complex (RISC) that mediates gene silencing via sequence-specific mRNA degradation or translational repression (Fig. **1a**).

Most mRNA fragments resulted from miRNA-mediated cleavage are degraded via general cellular mRNA degradation machinery. However, in some cases, the cleaved products are protected by SUPPRESSOR OF GENE SILENCING 3 (SGS3) and then converted to double-strand RNAs (dsRNAs) by RNA-dependent RNA polymerase 6 (RDR6). Subsequently, Dicer-like protein (DCL) recognizes these dsRNAs and chops them at a consistent interval into secondary siRNAs, which are designated as phased siRNAs (phasiRNAs) due to their continuous head-to-tail arrangement in phase to the miRNA cleavage site (Fig. **1a**) (Fei *et al.*, 2013). DCL4 is responsible for 21-nt phasiRNA generation. PhasiRNAs can amplify the gene silencing efficiency by *cis*-cleaving their precursor or *trans*-regulating other target genes simultaneously (Liu *et al.*, 2020). The miRNAs instigating phasiRNA production from their target genes, known as miRNA triggers, are usually 22-nt in length and commonly have uracil (U) as the first nucleotide in their sequence. PhasiRNAs have emerged as a major sRNA class in plants possessing diverse functions (reviewed in detail in Fei *et al.*, 2013; Liu *et al.*, 2020).

miR482/2118 (22nt) is widely present in seed plants and trigger phasiRNA biogenesis from its target genes (Fei *et al.*, 2013). miR482/2118 plays a critical role in various biological processes,

such as disease resistance and plant development, by targeting diverse genes, including both coding and noncoding genes, and generating profuse phasiRNAs, forming another layer of regulation to reinforce its silencing effect (Fei *et al.*, 2013; Liu *et al.*, 2020). In this review, we aim to provide a comprehensive view of the origin, evolutionary conservation, and functional diversification of the miR482/2118 superfamily, generating new insights into the biological significance of its functional diversification.

## Ⅱ. An overview of miR482/2118 Superfamily

miR482/2118 is a miRNA superfamily consisting of two founding miRNAs: miR482 and miR2118. The first miR482 was identified and cloned in *Populus trichocarpa* (*P. trichocarpa*) (Lu *et al.*, 2005), while miR2118 was later identified and specifically expressed in the reproductive tissues of rice and maize (Nobuta *et al.*, 2008; Johnson *et al.*, 2009). Initially, these miRNAs were thought to be completely distinct until 2012, when miR482s and miR2118s were combined into a superfamily due to their highly similar sequences and thereafter termed the miR482/2118 superfamily (Shivaprasad *et al.*, 2012). miR482/2118 are generated via a classical miRNA biosynthesis pathway (Fig. **1a**). According to the sequence alignment of miR482/2118 isoforms (Fig. **1b**), the "miR482-type" frequently starts with "UC", while the "miR2118-type" typically have a 2-nt shift backward. Because of this shift, these 22-nt miRNAs have a 20-nt overlapping region sequence conservation (Fig. **1b**). In many cases, both types of miRNA sequence can be generated from a single precursor, with the 2-nt shift the only difference. Similar to most 22-nt miRNAs derived from hairpin structures with an asymmetric bulge, miR482/2118 isoforms have a conserved bulge at the $10^{th}$ position according to the miR482-type sequence (Fig. **1b**).

Sequences of both the miR482 and miR2118 types are predominated by a "U" at the first nucleotide position, especially for the miR2118 type (Fig. **1b**), suggesting their association with AGO1 protein, which sorts miRNAs with a strong preference of a 5' terminal "U" (Kim, 2008). Precursors of miR482/2118 from different species share a substantial level of sequence

conservation in the lower stem region below the miRNA/miRNA* duplex, which maintains proper stem structure (Fig. **1c**), necessary for the recognition and base-to-loop processing by DCL1 (L. Song *et al.*, 2010; C. Song *et al.*, 2010; Werner *et al.*, 2010). Therefore, *MIR482/2118* copies are likely processed in a base-to-loop manner, like many other miRNAs such as miR172 and miR390 (Chorostecki *et al.*, 2017; Xia *et al.*, 2017). Regarding the subcellular location of this pathway, miR2118, *PHAS* precursors, and 21-nt reproductive phasiRNAs all displayed enrichment in membrane-bound polysomes (MBPs) in rice, and miR2118 and phasiRNA mediated cleavage can be detected in MBPs as well (Yang *et al.*, 2021), indicating this pathway is mainly associated with the rough endoplasmic reticulum, similar to other phasiRNA biogenesis pathways (Li *et al.*, 2016).

Like animal *MIR* genes, *MIR* clusters also exist in plants (Altuvia *et al.*, 2005; Xia *et al.*, 2015b). Many *MIR482/2118* have been reported to be present in a *MIR* cluster with their homologs or other related miRNAs. In the rice genome, 18 members of miR2118 are clustered at two genomic regions on chromosomes 4 and 11, encoding 12 different mature miR2118 isoforms (Johnson *et al.*, 2009). Four miR482 clusters encoding two miR482s each were found in the genome of litchi (*Litchi chinensis* Sonn.) (Ma *et al.*, 2018). In Solanaceae plants, miR482b-e are clustered on chromosome 6 in the progenitors of tomato and potato, while can-miR482f and species-specific can-miR482d are clustered on chromosome 4 in pepper but not in tomato and potato (Li *et al.*, 2011). These miR482/2118 clusters are potentially transcribed into mono-cistronic transcripts, then processed into distinct miRNAs, suggesting a co-regulation of these miR482/2118 isoforms at the transcription level.

## Ⅲ. Evolution and diversification of miR482/2118

Apart from the 2-nt shift, miR482 and miR2118 were initially designated as different miRNAs because of their distinct functions. miR482, identified initially in eudicot plants, was found to mainly target *NUCLEOTIDE BINDING SITE-LEUCINE-RICH REPEATS* (*NBS-LRRs*), which are essential components in disease resistance in plants (Zhai *et al.*, 2011; Shivaprasad *et al.*, 2012),

while miR2118, firstly identified in rice, targeted a large number of long non-coding transcripts (lncRNAs) with unknown function in the anther (Johnson *et al.*, 2009; Komiya *et al.*, 2014). More comprehensive analyses demonstrated an evolutionary divergence within miR482/2118, predominantly targeting hundreds of *NBS-LRR* genes with subsequent phasiRNA production in eudicots but numerous lncRNA genes of phasiRNA production in reproductive tissues of monocots (Fig. **2a**). Our examination of this miRNA family in spruce, a gymnosperm species, demonstrated that miR482/2118 could target both *NBS-LRR* genes, with phasiRNAs universally produced in vegetative tissues, and lncRNAs, inciting phasiRNA production specifically in reproductive tissues (male cones or female cones, Fig. **2a**) (Xia *et al.*, 2015a). Thus miR482/2118 has dual functions in gymnosperms, which seems to have been selectively retained in eudicots and monocot grasses (Fig. **2a**) ( Xia *et al.*, 2015a; Canto-pastor *et al.*, 2019). It is well-known that the target relationship of a conserved miRNA is always well-maintained during evolution, i.e., their target genes are similar among different plants. For instance, miR156 consistently targets *SPL* genes while miR172 targets *AP2* homologs (Chen, 2004; Wang & Wang, 2015). Therefore, this functional divergence of miR482/2118 targeting different genes among main plant lineages is rare in plants. Conceivably, it has resulted from the distinct diversification process of *NBS-LRR* between eudicots and monocots.

In plants, NBS-LRRs can be grouped into two major classes based on their N-terminal domains: TIR-NBS-LRRs (TNLs) and CC-NBS-LRRs (CNLs), with a Toll-interleukin-like receptor (TIR) homology domain or a coiled-coil (CC) motif-containing domain, respectively (Meyers *et al.*, 1999, 2003). TIR-NBS-LRRs first emerged in Chlorophyta and the other type CC-NBS-LRRs showed up in land plants, and all NBS-LRRs further expanded subsequently in gymnosperms (Fig. **2b**) (Zhang *et al.*, 2016; de Vries *et al.*, 2018a; Gao *et al.*, 2018). Following the origin of NBS-LRR, the most ancient miR482/2118 emerged in gymnosperm (Xia *et al.*, 2015a; de Vries *et al.*, 2018a) and then became widely present in all the angiosperms (Fig. **2b**). miR482/2118 is a relatively large *MIRNA* family, with more than 20 members in some species, such as the 24

isoforms encoded in the *Picea abies* genome (Xia *et al.*, 2015a). Although most plant species possess both the miR482-type and miR2118-type isoforms, the miR482-type expanded more widely in dicots while miR2118-type expanded in monocots (Fig. **2b**). The emergence of miR482/2118 in gymnosperms coincided with the expansion of *NBS-LRR* genes (Xia *et al.*, 2015a; de Vries *et al.*, 2018a), suggesting co-evolution between miR482/2118 and its targets *NBS-LRRs*, where the diversification of plant *NBS-LRR* genes directs the evolution of miRNAs that target them (Zhang *et al.*, 2016). Thus, during evolution, the appearance of *MIR482/2118* genes is correlated with the number of *NBS-LRR* genes in a species-specific manner (González *et al.*, 2015).

There are several hypotheses for *de novo* miRNA evolution, such as inverted duplication of target genes, inverted-repeat transposable elements (TE), and spliced-out introns (Budak & Akpinar, 2015; Baldrich *et al.*, 2018). Among these, the most well-accepted model for miRNA origination is via the inverted duplication of target genes proposed by Allen et al. (Allen *et al.*, 2005). Likewise, the miR482/2118 superfamily originated from its target *NBS-LRR* genes via inverted duplication, as shown in Fig. **2c** (Xia *et al.*, 2015a). At first, a spontaneous genomic inverted duplication of an *NBS-LRR* gene generated a perfect stem-loop-structure and acted as the backbone of "young miR482/2118". Over the course of evolution, these "young miRNA" accumulated mutations which disrupted the complementarity of its stem-loop structure, facilitating the recognition and accurate processing of *MIRNA* precursors by DCL1 into the miR482/2118 duplex, forming a canonical *MIRNA* locus (Allen *et al.*, 2005; Cui *et al.*, 2017; Baldrich *et al.*, 2018). Once the first miR482/2118 was born, several mechanisms were likely responsible for its evolutionary diversification. A series of duplication events led to the expansion of *MIRNA* genes and the target genes, such as *NBS-LRR*s (Baldrich *et al.*, 2018). This expansion is seen in the two *MIR2118* tandemly duplicated clusters in the rice genome. Transposable elements (TEs), which play a critical role in genetic diversity in plants, also contribute significantly to *MIR482/2118* proliferation, as well as loss or gain of *NBS-LRRs* prior to polyploidization as seen in cotton (Shen

*et al.*, 2020). Additionally, miR482/2118 can influence the evolution of its target genes, as seen in soybean, where the miRNA binding sites of target genes showed an evolutionary pattern different from the flanking sequences (Zhao *et al.*, 2015). In Solanaceae, more varied miR482/2118 sequences led to increased and highly dynamic targets, suggesting that this family acts as an evolutionary buffer for NBS-LRR sequence diversity (de Vries *et al.*, 2015).

Conversely, *NBS-LRR* genes also drive the evolution and diversification of miR482/2118. miR482/2118 targets *NBS-LRR* genes at the P-loop motif within the NB domain, a highly conserved and functionally important sequence especially for CNLs (Meyers *et al.*, 2003). Thus, miR482/2118 can target a large number of *NBS-LRRs*, mostly CNLs, in eudicots. The P-loop region encodes an amino acid sequence of "GMGG[V/I]G" (Shivaprasad *et al.*, 2012), which has unique complementarity to miR482/2118, a result of synergetic conservation and diversification of the miRNA with its targeting nucleotides. Large-scale comparative analysis revealed that the nucleotide diversity of the P-loop motif, especially the wobble position of the codons in the target site of miR482/2118, drives the diversification of this miRNA family, i.e., the sequences of miR482/2118 in seed plants show a higher degree of sequence variation at positions pairing with the wobble positions of their target sites (Fig. **2d**) (Zhang *et al.*, 2016). The periodic nucleotide diversity of miR482/2118 isoforms corresponding to the wobble position was more dramatic in eudicots and gymnosperms than in monocots (Zhang *et al.*, 2016). These distinct patterns of sequence diversity reflect the functional differences of miR482/2118 between eudicots/gymnosperms and monocots. These observations indicate that miR482/2118 has diversified to keep pace with their target *NBS-LRR* genes (Zhang *et al.*, 2016). Co-evolution between miR482/2118 and their targets also might be the primary force driving "duplicated *MIR482*" to evolve into other *NBS-LRR*-targeting miRNAs (Zhao *et al.*, 2012). The co-evolution and diversification of miR482/2118 and other related miRNAs is likely a strategy of plants to balance the costs and benefits of defense response.

## Ⅳ. Functions of miR482/2118

The miR482/2118 superfamily is widely present in seed plants and has evolved with diverse functions. Its ability to initiate phasiRNAs from target genes enhances and broadens its function in various biological pathways.

### 1. miR482/2118 in disease resistance

In eudicots, the primary function of miR482/2118 is to regulate *NBS-LRR* genes, a class of R-genes essential for plant defense. As shown in Fig. **3a**, the miR482/2118-targeted P-loop region is in the middle of the NBS domain of *NBS-LRRs*. By targeting conserved sequence motifs, miR482/2118 can regulate tens or hundreds of *NBS-LRR* genes in a single species. NBS-LRR proteins are an essential component of the plant immune system, which is composed of two layers of defense: pattern-triggered immunity (PTI) and effector-triggered immunity (ETI) (Fei *et al.*, 2016; Han, 2019). The zig-zag model of plant innate immunity can best explain the molecular basis for host-pathogen interactions (Fig. **3b**) (Jones & Dangl, 2006). PTI is a frontline of defense when membrane-associated pattern recognition receptors (PRRs) perceive pathogen-associated molecular patterns (PAMP). The pathogen deploys effectors to counteract with PTI. Pathogen-effectors (PE) activate effector-triggered susceptibility (ETS) followed by ETI signaling cascades. Compared with PTI, ETI is a rapid and high-amplitude output, in which evolved R-genes (NBS-LRRs) mediate intracellular perception of PEs and lead to hypersensitive responses.

miR482/2118 targets a large number of *NBS-LRR*s and triggers phasiRNA production, particularly in eudicots, and these ensuing phasiRNAs can act *in cis* or *trans* to reinforce the silencing effects of miR482/2118 on *NBS-LRRs* (Canto-pastor *et al.*, 2019), suggesting that miR482/2118 and subsequent phasiRNAs are integrative parts of the ETI in eudicots (Fig. **3b**) (Fei *et al.*, 2016). Thus the miR482/2118 family is a master regulator of disease resistance (Zhai *et al.*, 2011). Additionally, several long non-coding RNAs (lncRNAs) could also contribute to

miR482-mediated resistance (Song *et al.*, 2021). In tomato, miR2118b targets a lncRNA known as *TAS5*, which was formed by a rearrangement of several *CNLs* and *TNLs*, and thus *TAS5*-derived siRNAs can act *in trans* to regulate multiple *NBS-LRRs* (Canto-pastor *et al.*, 2019). Inactivation of miR2118b via short tandem target mimics (STTM) enhanced resistance to *P. syringae* and *P. infestans* (Canto-pastor *et al.*, 2019). Another tomato lncRNA, lncRNA23468, acted as an endogenous target mimic to decoying miR482b expression; silencing lncRNA23468 led to the increased accumulation of miR482b and decreased accumulation of *NBS-LRRs*, as well as reduced resistance to *P. infestans* in tomato (Jiang *et al.*, 2019).

The role of miR482/2118 in plant resistance has been confirmed in many plants in response to various pathogens. miR482/2118 can act as either a positive or negative regulator upon different pathogenic infections in different species. The mutant of miR472, a homologous counterpart of miR482 in Arabidopsis, was more resistant to *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto* DC3000) expressing a bacterial effector, whereas transgenic plants with *MIR472* overexpression were more susceptible to this bacterial strain (Boccara *et al.*, 2014). In soybean, transgenic expression of miR482 led to significant increase of nodule numbers, suggesting it may positively contribute to the nodulation formation in legume-rhizobial symbiosis (Li *et al.*, 2010). Consistent with this, miR482 was up-regulated in response to rhizobia bacteria *Bradyrhizobium japonicum* (*B. japonicum*) infection (Subramanian *et al.*, 2008). While in tomato, inactivation of miR482 and miR2118b via short tandem target mimics (STTM) enhanced resistance to *Pseudomonas syringae* (*Ps. synringae*) and *Phytophthora infestans* (*P. infestans*), suggesting a negative regulation of miR482/2118 on these bacterium and oomycete infection (Canto-pastor *et al.*, 2019). Upregulation of miR482/2118 early in the infection may increase susceptibility to *P. infestans* , and miR482a and miR482f are the candidate miRNAs important for the resistance response of tomato to *P. infestans* (de Vries *et al.*, 2018b). Under the infection of the fungus *Fusarium oxysporum* (*F. oxysporum*), which causes worldwide tomato vascular wilt disease, miR482/2118 showed varied responses in susceptible (Moneymaker) and resistant (Motelle) tomato cultivars. miR482f showed

a reduction in both cultivars (Ouyang *et al.*, 2014), while miR482e was down-regulated in Motelle but up-regulated in Moneymaker (Ji *et al.*, 2018). Similarly, this opposite reaction of miR482/2118 could also be observed in plants invaded by viruses. In tomato, miR482 was up-regulated in co-infection of virus tomato yellow leaf curl China virus (TYLCCNV) and a betasatellite (TYLCCNB) in *Nicotiana benthamiana* (Xiao *et al.*, 2014) while down-regulated after groundnut bud necrosis virus (GBNV) infection in cowpea (Permar *et al.*, 2014). These opposite responses of miR482/2118 isoforms indicated that miR482/2118 members might be under different regulatory mechanisms in response to different pathogens.

## 2. miR482/2118 in plant development

In addition to its role in plant defense, the other primary function of the miR482/2118 family is plant reproductive development. In grasses, there are abundant 21-*PHAS* loci, which can generate 21-nt phasiRNAs. These phasiRNAs are triggered by miR2118, preferentially expressed in anther at the premeiotic stage (Johnson *et al.*, 2009; Song *et al.*, 2012; Zhai *et al.*, 2015; Fan *et al.*, 2016). The reproductive 21-nt phasiRNAs showed wide variation in abundance because they can act *in cis* to shift the pattern of phasiRNA production from miR2118 precursor (Tamim *et al.*, 2018; Tian *et al.*, 2021) and also *in trans* to target other genes, resulting in plant fertility regulation (Jiang *et al.*, 2020; Zhang *et al.*, 2020).

In rice, different 21-nt phasiRNAs produced from miR2118-targeted lncRNAs may exert their functions via different machinery. MEIOSIS ARRESTED AT LEPTOTENE1 (MEL1), a rice AGO protein homologous to *Arabidopsis* AGO5, associates with 21-nt phasiRNAs with a cytosine ("C") at the 5' terminal. MEL1 has a specific role in the development of pre-meiotic germ cells and the progression of meiosis (Komiya *et al.*, 2014). Recently, a study using low-input degradome sequencing in germ cells uncovered that the MEL1-associated "C"-initial phasiRNAs could target hundreds of protein-coding genes. These genes are mostly enriched in the pathways associated with carbohydrates biosynthesis and metabolism, which are critical for male germ cell

development (Fig. **3c**) (Jiang *et al.*, 2020). Another study demonstrated that thousands of genes, which were significantly enriched in adenyl ribonucleotide binding, kinase activity, and hydrolase activity, were cleaved by reproductive phasiRNAs (Zhang *et al.*, 2020). It was proposed that these miR2118-mediated reproductive phasiRNAs were essential for *trans* RNA cleavage and mRNA elimination (e.g., ATP binding genes and transcription related genes) during meiotic prophase I, ensuring proper meiotic progression (Zhang *et al.*, 2020). Interestingly, a recent study reveals that miR2118-dependent 21-nt phasiRNAs in the anther wall are U-rich at their 5' terminal, which is distinct from those in germ cells (Araki *et al.*, 2020). A large CRISPR-induced deletion of the miR2118 cluster in chromosome 4 resulted in the loss of 14 out of the 18 *MIR2118* loci in the rice genome and a subsequent reduction of "U"-initial phasiRNAs. The rice CRISPR mutant showed severe defects in inner anther wall layers and tapetum differentiation. This is because these phasiRNAs affect anther-wall specific transcription factors (TFs) such as TAPETUM DEGENERATION RETARDATION1 (TDR1), TDR INTERACTING PROTEIN2 (TIP2), and AGO1b/AGO1d (Araki *et al.*, 2020) (Fig. **3c**). How the initial nucleotide differences lead to different roles for these phasiRNA remains unclear. Likely, they are transported to different cell layers, possibly by being loaded to different AGO proteins, after their synthesis (Zhai *et al.*, 2015).

Rice varieties with photoperiod-sensitive and thermo-sensitive genic male sterility (P/TGMS) are essential for the development of two-line hybrid rice, which has made a tremendous contribution to rice production in China (Ding *et al.*, 2012; Zhou *et al.*, 2012; Fan *et al.*, 2016). There are two gene loci conferring PGMS in rice, *PMS1(T)* and *PMS3*. Interestingly, both loci encode lncRNAs targeted by miR2118 to generate secondary phasiRNAs (Ding *et al.*, 2012; Zhou *et al.*, 2012; Fan *et al.*, 2016). Moreover, functional mutation of male fertility is caused by SNPs at both loci (G → T change in *PMS1T* and a C → G change in *PMS3*), and both SNPs are in the 2$^{nd}$ phase cycle of phasiRNA (Fig. **3c**) (Ding *et al.*, 2012; Zhou *et al.*, 2012; Fan *et al.*, 2016). However, there are differences between these two loci; sterility is of incomplete dominance for *Pms1*, whereas recessive for *Pms3*. A sufficient amount of *PMS3* transcript is required for fertility in long days,

whereas high expression of *PMS1T* and its ensuing phasiRNAs results in sterility (Ding *et al.*, 2012; Zhou *et al.*, 2012; Fan *et al.*, 2016). The fertility of PGMS rice under long-day conditions is governed by a lncRNA, *PMS1T,* and is expressed specifically in panicle development. *PMS1T* is a typical miR2118 target with the production of 21-nt phasiRNAs, which may function as siRNAs to act on some unknown downstream targets to affect male fertility. However, how a single nucleotide polymorphism (SNP) near the miR2118 target site causes differential accumulation of phasiRNAs, affecting their regulation of downstream target genes, and eventually, the change of fertility remains obscure. Although a recent study by Jiang et al. (2020) seems to validate that these lncRNA-derived phasiRNAs exert function via the mediation of mRNA cleavage, no highly confident target genes have been identified for *Pms1*- and *Pms3*-derived phasiRNAs, especially these phasiRNAs harboring the functional SNPs. Therefore, it is still unclear whether other regulatory mechanisms are involved in this process, like the RNA-directed DNA methylation suggested in *PMS3* (Ding *et al.*, 2012). miR2118-mediated premeiotic reproductive phasiRNAs are not only broadly present in many monocots, such as maize, rice, *Asparagus Officinalis, Lilium maculatum,* and *Hemerocallis lilioasphodelus* (Kakrana *et al.*, 2018), but also in some eudicots such as flax, strawberry, and columbine, although via different miRNA triggers (Pokhrel *et al.*, 2021), implying that this miR2118-related premeiotic phasiRNA regulation may be conserved in divergent modern flowering plants (Zhai *et al.*, 2015; Fan *et al.*, 2016; Kakrana *et al.*, 2018).

## 3. miR482/2118 in abiotic stress and other process

In addition to its main roles in disease resistance and plant development, as aforementioned, the dynamic evolution of the miR482/2118 superfamily gives it a remarkable ability to obtain new target genes, which perform diverse functions. Novel target genes have been identified for miR482/2118 in a few species. In litchi, in addition to the well-defined targets, many additional targets were identified, including those genes showing homolog to F-box protein, phosphoglycerate kinase, S-adenosylmethionine carrier, etc., and a few of these targets are likely transcribed into non-coding transcripts, which generate secondary phasiRNAs (Ma *et al.*, 2018).

miR2118 also affected fruit storability via targeting energy metabolism-related genes such as PGK and $Ca^{2+}$-ATPase in litchi (Tang *et al.*, 2020). A similar expansion of target genes was also discovered in strawberry, in which miR482 targets genes encoding ribonuclease H protein, transparent testa, and cyclic phosphodiesterase (Xia *et al.*, 2015b). In *P. trichocarpa*, miR482 targeted four genes encoding UDP-glucosyltransferases, suggesting a role miR482 in stress adaption as well (Shuai *et al.*, 2013).

## Ⅴ. Perspective

Since the first discovery of miR482/2118 16 years ago, we have gained a much deeper understanding of its origin, biogenesis, evolution, and functional diversity. Compared to other conserved miRNAs, the miR482/2118 family members have greater diversity on miRNA sequences and target genes. Thus, we still have a few unanswered questions worthy of further investigation.

(1) What is the upstream regulatory network of the miR482/2118 family? Are the different *MIR482/2118* genes of functional difference?

*MIR482/2118* usually consists of 8~20 members, with a few of them producing miRNAs of the same sequence in a species. Conceivably, it is similar to a protein-coding gene family where multiple members usually confer redundant as well as specific functions. As different miRNAs derived from different *MIRNA* genes are of identical or highly similar sequences, it is reasonable to believe that different miR482/2118s are of redundant function. But in an alternative scenario, different *MIRNA* genes could be under different regulatory mechanisms due to the fact that they possess different promoters and thus may be able to respond to different developmental signals. How miR482/2118s convey member-specific function is an interesting question to investigate. Likely, minor sequence differences among miRNA members can incite distinct preferences of target genes. For instance, although miR482/2118s can target many *NBS-LRR* genes, members of miR482/2118 probably compete for their target genes. As NBS-LRR proteins are sensors to

virulence effectors from specific pathogens and NBS-LRR and effector protein are always in a gene-for-gene manner, preferential targeting of miR482/2118 may link specific members to given pathogens. This is supported by the fact that certain pathogen infections activated different miR482/2118 members. Systematic construction of the expression map of miR482/2118s and their target genes in different developmental stages or under various stress conditions will be helpful to dissect the routes of specific miR482/2118 members to peculiar target genes in a given condition.

(2) What is the function of so many non-coding *PHAS* loci targeted by miR2118 in grasses (e.g., >2000 loci in rice genome)? Are their functions distinct or similar? How do they perform their functions?

There are hundreds or even thousands of non-coding 21-*PHAS* loci targeted by miR2118 in a single grass species, and sequences of these *PHAS* loci often lack many similarity besides the miRNA target site. Therefore, there are a large amount of phasiRNAs with distinct sequences produced at the premeiotic stage of anther. Based on what we know, phasiRNAs generated from *PHAS* loci can function via directing mRNA cleavage of target genes by reverse complementary sequence paring. Given the large quantity of reproductive phasiRNAs, computational prediction shows that almost all of the protein-coding genes (e.g., rice) have the potential to be targeted by a reproductive phasiRNA. Although studies have proven that reproductive phasiRNAs have the ability to direct mRNA cleavage, it is hard to believe that all the phasiRNAs are involved in target cleavage. Even using highly sensitive degradome sequencing, only ~250 genes were proven to be targeted by 21-nt phasiRNAs (Jiang *et al.*, 2020).

Another intriguing feature of these reproductive *PHAS* loci is their great sequence diversity. There is little sequence similarity among different *PHAS* loci within a species, and significant sequence divergence is represented in homologous *PHAS* loci among plant species from the same genus (e.g., *Oryza* species) (Tian *et al.*, 2021). The broad presence and conserved anther-specific expression of reproductive phasiRNAs in grasses and other flowering plants (Pokhrel *et al.*, 2021) indicate that these *PHAS* and phasiRNAs are of similar or conserved function in reproduction.

This is in sharp contrast to their great sequence difference, an implication of divergent target genes. In other words, how are different target genes regulated by diverse phasiRNAs convey a similar or conserved function among different species? Three possibilities may exist:

a) The role of these divergent and highly abundant phasiRNAs is to quickly turn down or turn off the expression of most genes to allow quick reprogramming of the premeiotic anther transcriptome. This scenario is a "scattergun" strategy, in which large quantities of diverse phasiRNAs universally cut off the majority of genes to help reprogram cell development. In this case, there is no selection on which genes of which pathways will be silenced. Therefore, there is no need to maintain sequence conservation of *PHAS* loci and phasiRNAs.

b) Another possibility points to that the main function of reproductive phasiRNA is not the *trans* activity, but the *cis* action of phasiRNA, i.e., the backward targeting of phasiRNAs to the cognate *PHAS* mRNA (mRNA of phasiRNA precursor). This targeting would not necessarily cause mRNA cleavage but would mediate the loading of certain protein complexes to the *PHAS* loci in the genome, where the *PHAS* mRNAs play a role as a scaffold for protein complex localization. In rice, reproductive *PHAS* loci are of genomic position, not sequence conservation among divergent *Oryza* species (Tian *et al.*, 2021), suggesting their genomic position is possibly more functionally important than their sequences. Therefore, the *cis*-activity of phasiRNAs may be important, e.g., for the chromosome condensation at early meiotic stages, which is shown in mel1 (ago5) mutant in rice (Nonomura *et al.*, 2007).

c) In addition, there may be other unknown components involved in the pathway. The transcription of many *PHAS* loci is complicated, with different transcript variants, which not only have distinct transcription start or polyadenylation sites but also can be transcribed from different strands. For example, the *PMS3* locus in rice, essential for photoperiod-sensitive male sterility, has mRNA transcripts transcribed from both DNA strands, and some have alternative splicing (Ding *et al.*, 2012). Are these alternative transcripts of functional importance, especially for the antisense transcripts, reverse complementary to the functional *LDMAR* transcript? It is conceivable to

believe the antisense transcripts are functional because they can form double-strands with the *LDMAR* transcript and also have the potential to be targeted by phasiRNAs.

(3) How are the strict temporal- and spatial-specific expression of miR2118 and its target *PHAS* genes regulated in grasses, via a transcriptional regulation by transcription factors or other unknown mechanisms?

In grasses, miR2118, its target *PHAS* genes, and ensuing phasiRNAs are expressed explicitly in the pre-meiotic anther stage. If this strict temporal- or spatial- expression is achieved by the regulation of a common transcription factor, it would make sense for the promoters of these *MIRNA* or *PHAS* genes to have a shared binding site for a TF. But this is unlikely because: (1) based on broad sequence comparative analysis of the promoter sequences of the *PHAS* loci, no commonly conserved DNA motifs can be identified; (2) in some *PHAS* clusters, two adjacent loci are very close to each other, with a <500 bp distance, suggesting they have no or very short promoters. Therefore, *PHAS* loci may have different transcriptional regulation mechanisms. Given that the miR482/2118 target site is the only region highly conserved and under strong selection in the sequences of *PHAS* genes, if there is a common mechanism controlling their transcription, the target site could probably be the shared steering motif acting as a binding site for a TF. Alternatively, it could be a scaffolding site for other transcription machinery, for instance, the AGO proteins, which can also regulate transcription in an sRNA-dependent way (Liu *et al.*, 2018). If a specific TF can recognize the miR482/2118 target site (DNA) and trigger the simultaneous transcription of all the *PHAS* loci, the DNA sequence of the target site can be used as bait to identify potential TFs. If any AGO proteins are involved in the transcription, systematic analysis of all *AGO* genes in rice might be applicable, especially AGO1 homologs and AGO5 (MEL1), which are associated with miR2118 and 21-nt reproductive phasiRNA, respectively.

(4) How are the species-specific target genes gained or lost by miR482/2118? What is the underlying mechanism for the fast evolution of novel target genes of miR482/2118 members? Are their target genes connected or unrelated in terms of biological function?

Compared to other well-known miRNAs, one of the striking features of the miR482/2118 is that, other than well-known targets like NBS-LRRs, it has more lineage- or species-specific target genes. This can be explained by: (1) in contrast to other miRNAs, miR482/2118 family has more members producing more miRNA isoforms with greater sequence divergence (and allowing 2-nt shift), enabling it to have a higher ability to capture new target genes during evolution; (2) the large quantity of target genes, especially, the NBS-LRR family (usually with >100 members in a single species) which endows a high possibility of sequence shuffle or exchange, and quick evolution of target genes. The relatively fast diversification of *NBS-LRR* genes in the host-pathogen arms race may contribute to the quick acquirement of new target genes for miR482/2118 as well. These newly acquired target genes, could be involved in similar biological processes (target genes evolved from homologous genes or of similar expression profile) or completely unrelated pathways if their target genes evolved from random sequence exchange.

(5) How are the miR2118-target non-coding *PHAS* loci evolved? Are they derived from the pseudogenization of miR2118 targeted coding genes? How are they amplified in the grass genomes? Are they under certain selection pressure?

*MIR482/2118* gene firstly appeared in seed plants through the inverted duplication of its targeted *NBS-LRR* genes, and miR2118-initiated reproductive phasiRNAs are also present in Norway spruce. It is unlikely that non-coding *PHAS* loci are derived from pseudogenization of miR2118-target coding genes because reproductive *PHAS* loci in Norway spruce share little sequence similarity to *NBS-LRR* genes. Alternatively, miR482/2118 could have had the ability to target non-coding reproductive *PHAS* genes once it appeared, and these *PHAS* genes may have evolved earlier or simultaneously with *MIR482/2118s*. There are hundreds of or even thousands of reproductive *PHAS* genes in a single species, larger than most protein-coding genes, implying that

these *PHAS* genes have a greater amplification rate than coding genes, with many of them forming tandem or segmental duplications (Tian *et al.*, 2021). The presence of superclusters of ~100 *PHAS* loci suggests these genes utilize a different expansion strategy than coding genes. Moreover, genomic position conservation, not sequence conservation, of these non-coding *PHAS* loci is another indication of their difference from protein coding genes, and likely they are under different selection during evolution.

**Acknowledgements**

We thank other members of the Xia lab for helpful discussion and comments on the manuscript. We are grateful to Dr. Blake C. Meyers from the Donald Danforth Plant Science Center (USA) for his critical reading and suggestions. We apologize to authors whose work could not be included in this review, due to space limits. This work is supported by the National Science Foundation of China (#31872063), and the Special Support Program of Guangdong Province (#2019TX05N193).

**Author contributions**

YZ, MW, and RX designed the work, collected data, prepared figures, and wrote the review with input from all authors; ZZ, JX, and CC collected related information and prepared figures; YL, JZ and RX discussed and designed the structure of the review. YZ and MW contributed equally to this work.

## References

**Allen E, Xie Z, Gustafson AM, Carrington JC**. **2005**. microRNA-directed phasing during *trans*-acting siRNA biogenesis in plants. *Cell* **121**: 207–221.

**Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, Brownstein MJ, Tuschl T, Margalit H**. **2005**. Clustering and conservation patterns of human microRNAs. *Nucleic Acids Research* **33**: 2697–2706.

**Araki S, Le NT, Koizumi K, Villar-Briones A, Nonomura KI, Endo M, Inoue H, Saze H, Komiya R**. **2020**. miR2118-dependent U-rich phasiRNA production in rice anther wall development. *Nature Communications* **11**: 3115.

**Baldrich P, Beric A, Meyers BC**. **2018**. Despacito: the slow evolutionary changes in plant microRNAs. *Current Opinion in Plant Biology* **42**: 16–22.

**Boccara M, Sarazin A, Thie´beauld O, Jay F, Voinnet O, Navarro L, Colot V**. **2014**. The Arabidopsis *miR472-RDR6* silencing pathway modulates PAMP- and effector-triggered immunity through the post-transcriptional control of disease resistance genes. *PLoS Pathogens* **10**: e1003883.

**Budak H, Akpinar BA**. **2015**. Plant miRNAs: biogenesis, organization and origins. *Functional and Integrative Genomics* **15**: 523–531.

**Canto-pastor A, Santos BAMC, Valli AA, Summers W, Schornack S, Baulcombe DC**. **2019**. Enhanced resistance to bacterial and oomycete pathogens by short tandem target mimic RNAs in tomato. *Proceedings of the National Academy of Sciences, USA* **116**: 2755–2760.

**Chen X**. **2004**. A microRNA as a translational repressor of *APETALA2* in *Arabidopsis* flower development. *Science* **303**: 2022–2025.

**Chen C, Zeng Z, Liu Z, Xia R**. **2018**. Small RNAs, emerging regulators critical for the development of horticultural traits. *Horticulture Research* **5**: 63.

**Chorostecki U, Moro B, Rojas AML, Debernardi JM, Schapire AL, Notredame C, Palatnik JF**. **2017**. Evolutionary footprints reveal insights into plant microRNA biogenesis. *Plant Cell* **29**: 1248–1261.

**Cui J, You C, Chen X**. **2017**. The evolution of microRNAs in plants. *Current Opinion in Plant Biology* **35**: 61–67.

**D'Ario M, Griffiths-Jones S, Kim M**. **2017**. Small RNAs: big impact on plant development. *Trends in Plant Science* **22**: 1056–1068.

**Ding J, Lu Q, Ouyang Y, Mao H, Zhang P, Yao J, Xu C, Li X, Xiao J, Zhang Q**. **2012**. A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proceedings of the National Academy of Sciences, USA* **109**: 2654–2659.

**Fan Y, Yang J, Mathioni SM, Yu J, Shen J, Yang X, Wang L, Zhang Q, Cai Z, Xu C, *et al.* 2016**. *PMS1T*, producing phased small-interfering RNAs, regulates photoperiod-sensitive male sterility in rice. *Proceedings of the National Academy of Sciences, USA* **113**: 15144–15149.

**Fei Q, Xia R, Meyers BC**. **2013**. Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *Plant Cell* **25**: 2400–2415.

**Fei Q, Zhang Y, Xia R, Meyers BC**. **2016**. Small RNAs add zing to the zig-zag-zig model of plant defenses. *Molecular Plant-Microbe Interactions* **29**: 165–169.

**Gao Y, Wang W, Zhang T, Gong Z, Zhao H, Han GZ**. **2018**. Out of water: the origin and early diversification of plant R-genes. *Plant Physiology* **177**: 82–89.

**González VM, Müller S, Baulcombe D, Puigdomènech P**. **2015**. Evolution of *NBS-LRR* gene copies among dicot plants and its regulation by members of the miR482/2118 superfamily of miRNAs. *Molecular Plant* **8**: 329–331.

**Han G-Z**. **2019**. Origin and evolution of the plant immune system. *New Phytologist* **222**: 70–83.

**Ji H-M, Zhao M, Gao Y, Cao X-X, Mao H-Y, Zhou Y, Fan W-Y, Borkovich KA, Ouyang S-Q, Liu P**. **2018**. *FRG3*, a target of slmiR482e-3p, provides resistance against the fungal pathogen *Fusarium oxysporum* in tomato. *Frontiers in Plant Science* **9**: 26.

**Jiang N, Cui J, Shi Y, Yang G, Zhou X, Hou X, Meng J, Luan Y**. **2019**. Tomato lncRNA23468 functions as a competing endogenous RNA to modulate *NBS-LRR* genes by decoying miR482b in the tomato-*Phytophthora infestans* interaction. *Horticulture Research* **6**: 28.

**Jiang P, Lian B, Liu C, Fu Z, Shen Y, Cheng Z, Qi Y**. **2020**. 21-nt phasiRNAs direct target mRNA cleavage in rice male germ cells. *Nature Communications* **11**: 5191.

**Johnson C, Kasprzewska A, Tennessen K, Fernandes J, Nan G, Walbot V, Sundaresan V, Vance V, Bowman LH**. **2009**. Clusters and superclusters of phased small RNAs in the developing inflorescence of rice. *Genome Research* **19**: 1429–1440.

**Jones JDG, Dangl JL**. **2006**. The plant immune system. *Nature* **444**: 323–329.

**Kakrana A, Mathioni SM, Huang K, Hammond R, Vandivier L, Patel P, Arikit S, Shevchenko O, Harkess AE, Kingham B, *et al.* 2018**. Plant 24-nt reproductive phasiRNAs from intramolecular duplex mRNAs in diverse monocots. *Genome Research* **28**: 1333–1344.

**Kim VN**. **2008**. Sorting out small RNAs. *Cell* **133**: 25–26.

**Komiya R, Ohyanagi H, Niihama M, Watanabe T, Nakano M, Kurata N, Nonomura K-I**. **2014**. Rice germline-specific Argonaute MEL1 protein binds to phasiRNAs generated from more than 700 lincRNAs. *The Plant Journal* **78**: 385–397.

**Li H, Deng Y, Wu T, Subramanian S, Yu O**. **2010**. Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. *Plant Physiology* **153**: 1759–1770.

**Li S, Le B, Ma X, Li S, You C, Yu Y, Zhang B, Liu L, Gao L, Shi T, *et al.* 2016**. Biogenesis of phased siRNAs on membrane-bound polysomes in Arabidopsis. *eLife* **5**: e22750.

**Li F, Pignatta D, Bendix C, Brunkard JO, Cohn MM, Tung J, Sun H**. **2011**. MicroRNA regulation of plant innate immune receptors. *Proceedings of the National Academy of Sciences, USA* **109**: 1790–1795.

**Liu Y, Teng C, Xia R, Meyers BC**. **2020**. Phased secondary small interfering RNAs (phasiRNAs) in plants: their biogenesis, genic sources, and roles in stress responses, development, and reproduction. *Plant Cell* **32**: 3059–3080.

**Liu C, Xin Y, Xu L, Cai Z, Xue Y, Liu Y, Xie D, Liu Y, Qi Y**. **2018**. Arabidopsis ARGONAUTE 1 binds chromatin to promote gene transcription in response to hormones and stresses. *Developmental Cell* **44**: 1–14.

**Lu S, Sun Y-H, Shi R, Clark C, Li L, Chiang VL**. **2005**. Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. *Plant Cell* **17**: 2186–2203.

**Ma W, Chen C, Liu Y, Zeng M, Meyers BC, Li J, Xia R**. **2018**. Coupling of microRNA-directed phased small interfering RNA generation from long noncoding genes with alternative splicing and alternative polyadenylation in small RNA-mediated gene silencing. *New Phytologist* **217**: 1535–1550.

**Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND**. **1999**. Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *The Plant Journal* **20**: 317–332.

**Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW**. **2003**. Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *Plant Cell* **15**: 809–834.

**Nobuta K, Lu C, Shrivastava R, Pillay M, De Paoli E, Accerbi M, Arteaga-Vazquez M, Sidorenko L, Jeong DH, Yen Y, *et al.* 2008**. Distinct size distribution of endogenous siRNAs in maize: Evidence from deep sequencing in the *mop1-1* mutant. *Proceedings of the National Academy of Sciences, USA* **105**: 14958–14963.

**Nonomura KI, Morohoshi A, Nakano M, Eiguchi M, Miyao A, Hirochika H, Kurataa N**. **2007**. A germ cell-specific gene of the ARGONAUTE family is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice. *Plant Cell* **19**: 2583–2594.

**Ouyang S, Park G, Atamian HS, Han CS, Stajich JE, Kaloshian I, Borkovich KA**. **2014**. microRNAs suppress NB domain genes in tomato that confer resistance to *Fusarium oxysporum*. *PLoS Pathogens* **10**: e1004464.

**Permar V, Singh A, Pandey V, Alatar AA, Faisal M, Jain RK, Praveen S**. **2014**. Tospo viral infection instigates necrosis and premature senescence by micro RNA controlled programmed cell death in *Vigna unguiculata*. *Physiological and Molecular Plant Pathology* **88**: 77–84.

**Pokhrel S, Huang K, Bélanger S, Zhan J, Caplan JL, Kramer EM, Meyers BC**. **2021**. Pre-meiotic 21-nucleotide reproductive phasiRNAs emerged in seed plants and diversified in flowering plants. *Nature Communications* **12**: 4941.

**Shen E, Chen T, Zhu X, Fan L, Sun J, Llewellyn DJ, Wilson I, Zhu Q-H**. **2020**. Expansion of *MIR482/2118* by a class-II transposable element in cotton. *Plant Journal* **103**: 2084–2099.

**Shivaprasad P V, Chen H, Patel K, Bond DM, Santos BACM, Baulcombe DC**. **2012**. A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. *Plant Cell* **24**: 859–874.

**Shuai P, Liang D, Zhang Z, Yin W, Xia X**. **2013**. Identification of drought-responsive and novel *Populus trichocarpa* microRNAs by high-throughput sequencing and their targets using degradome analysis. *BMC Genomics* **14**: 233.

**Song L, Axtell MJ, Fedoroff N V. 2010a**. RNA secondary structural determinants of miRNA precursor processing in *Arabidopsis*. *Current Biology* **20**: 37–41.

**Song L, Fang Y, Chen L, Wang J, Chen X**. **2021**. Role of non-coding RNAs in plant immunity. *Plant Communications* **2**: 100180.

**Song C, Jia Q, Fang J, Li F, Wang C, Zhang Z**. **2010b**. Computational identification of citrus microRNAs and target analysis in citrus expressed sequence tags. *Plant Biology* **12**: 927–934.

**Song X, Li P, Zhai J, Zhou M, Ma L, Liu B, Jeong D**. **2012**. Roles of DCL4 and DCL3b in rice phased small RNA biogenesis. *The plant journal* **69**: 462–474.

**Subramanian S, Fu Y, Sunkar R, Barbazuk WB, Zhu JK, Yu O**. **2008**. Novel and nodulation-regulated microRNAs in soybean roots. *BMC Genomics* **9**: 160.

**Tamim S, Cai Z, Mathioni SM, Zhai J, Teng C, Zhang Q, Meyers BC**. **2018**. *Cis*-directed cleavage and nonstoichiometric abundances of 21-nucleotide reproductive phased small interfering RNAs in grasses. *New Phytologist* **220**: 865–877.

**Tang R fang, Zhou Y jie, Chen Z suzhi, Zeng J, Huang H, Jiang Y ming, Qu H xia, Zhu H**. **2020**. Involvement of miRNA-mediated anthocyanin and energy metabolism in the storability of litchi fruit. *Postharvest Biology and Technology* **165**: 111200.

**Tian P, Zhang X, Xia R, Liu Y, Wang M, Li B, Liu T, Shi J, Wing RA, Meyers BC, *et al.* 2021**. Evolution and diversification of reproductive phased small interfering RNAs in *Oryza* species. *New Phytologist* **229**: 2970–2983.

**de Vries S, Kloesges T, Rose LE**. **2015**. Evolutionarily dynamic, but robust, targeting of resistance genes by the miR482/2118 gene family in the solanaceae. *Genome Biology and Evolution* **7**: 3307–3321.

**de Vries S, de Vries J, von Dahlen JK, Gould SB, Archibald JM, Rose LE, Slamovits CH**. **2018a**. On plant defense signaling networks and early land plant evolution. *Communicative and Integrative Biology* **11**: 1–14.

**de Vries S, Kukuk A, von Dahlen JK, Schnake A, Kloesges T, Rose LE**. **2018b**. Expression profiling across wild and cultivated tomatoes supports the relevance of early miR482/2118 suppression for *Phytophthora* resistance. *Proceedings of the Royal Society B: Biological Sciences* **285**: 20172560.

**Wang H, Wang H**. **2015**. The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits. *Molecular Plant* **8**: 677–688.

**Werner S, Wollmann H, Schneeberger K, Weigel D**. **2010**. Structure determinants for accurate processing of miR172a in *Arabidopsis thaliana*. *Current Biology* **20**: 42–48.

**Xia R, Xu J, Arikit S, Meyers BC**. **2015a**. Extensive families of miRNAs and *PHAS* loci in Norway spruce demonstrate the origins of complex phasiRNA networks in seed plants. *Molecular Biology and Evolution* **32**: 2905–2918.

**Xia R, Xu J, Meyers BC**. **2017**. The emergence, evolution, and diversification of the miR390-*TAS3*-*ARF* pathway in land plants. *Plant Cell* **29**: 1232–1247.

**Xia R, Ye S, Liu Z, Meyers BC, Liu Z**. **2015b**. Novel and recently evolved microRNA clusters regulate expansive *F-BOX* gene networks through phased small interfering RNAs in wild diploid strawberry. *Plant Physiology* **169**: 594–610.

**Xiao B, Yang X, Ye C, Liu Y, Yan C, Wang Y, Lu X, Li Y, Fan L**. **2014**. A diverse set of miRNAs responsive to begomovirus-associated betasatellite in *Nicotiana benthamiana*. *BMC Plant Biology* **14**: 60.

**Yang X, You C, Wang X, Gao L, Mo B, Liu L, Chen X**. **2021**. Widespread occurrence of microRNA-mediated target cleavage on membrane-bound polysomes. *Genome Biology* **22**: 15.

**Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R, Chen X**. **2005**. Methylation as a crucial step in plant microRNA biogenesis. *Science* **307**: 932–935.

**Zhai J, Jeong D, Paoli E De, Park S, Rosen BD, Yan Z, Kitto SL, Grusak MA, Jackson SA, Li Y, *et al.* 2011**. MicroRNAs as master regulators of the plant *NB-LRR* defense gene family via the production of phased, *trans*-acting siRNAs. *GENES & DEVELOPMENT* **25**: 2540–2553.

**Zhai J, Zhang H, Arikit S, Huang K, Nan G-L, Walbot V, Meyers BC**. **2015**. Spatiotemporally dynamic, cell-type-dependent premeiotic and meiotic phasiRNAs in maize anthers. *Proceedings of the National Academy of Sciences* **112**: 3146–3151.

**Zhang YC, Lei MQ, Zhou YF, Yang YW, Lian JP, Yu Y, Feng YZ, Zhou KR, He RR, He H, *et al.* 2020**. Reproductive phasiRNAs regulate reprogramming of gene expression and meiotic progression in rice. *Nature Communications* **11**: 6031.

**Zhang Y, Xia R, Kuang H, Meyers BC**. **2016**. The diversification of plant *NBS-LRR* defense genes directs the evolution of microRNAs that target them. *Molecular Biology and Evolution* **33**: 2692–2705.

**Zhao J, Diao S, Zhang B, Niu B, Wang Q, Wan X, Luo Y-Q**. **2012**. Phylogenetic analysis and molecular evolution patterns in the *MIR482-MIR1448* polycistron of *Populus* L. *PLoS ONE* **7**: e47811.

**Zhao M, Meyers BC, Cai C, Xu W, Ma J**. **2015**. Evolutionary patterns and coevolutionary consequences of *MIRNA* genes and microRNA targets triggered by multiple mechanisms of genomic duplications in soybean. *Plant Cell* **27**: 546–562.

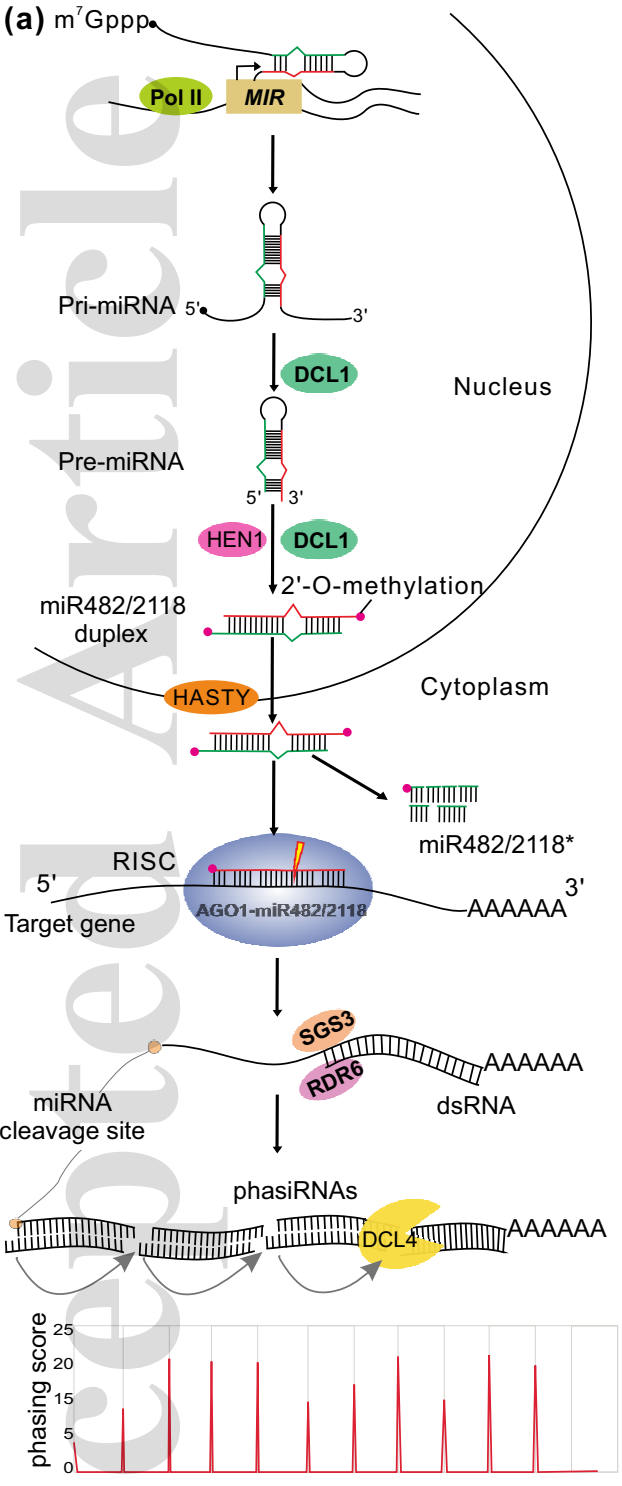
**Zhou H, Liu Q, Li J, Jiang D, Zhou L, Wu P, Lu S, Li F, Zhu L, Liu Z, *et al.* 2012**. Photoperiod- and thermo-sensitive genic male sterility in rice are caused by a point mutation in a novel noncoding RNA that produces a small RNA. *Cell Research* **22**: 649–660.

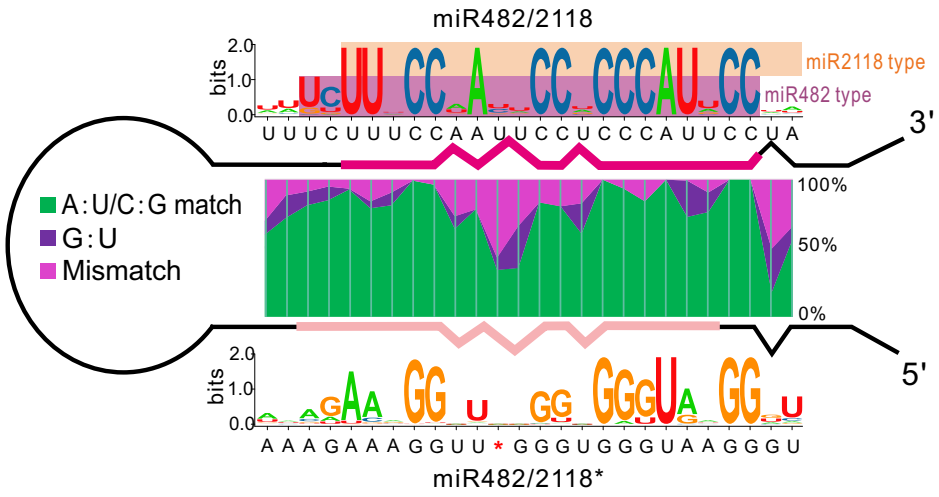
**Figure legends**

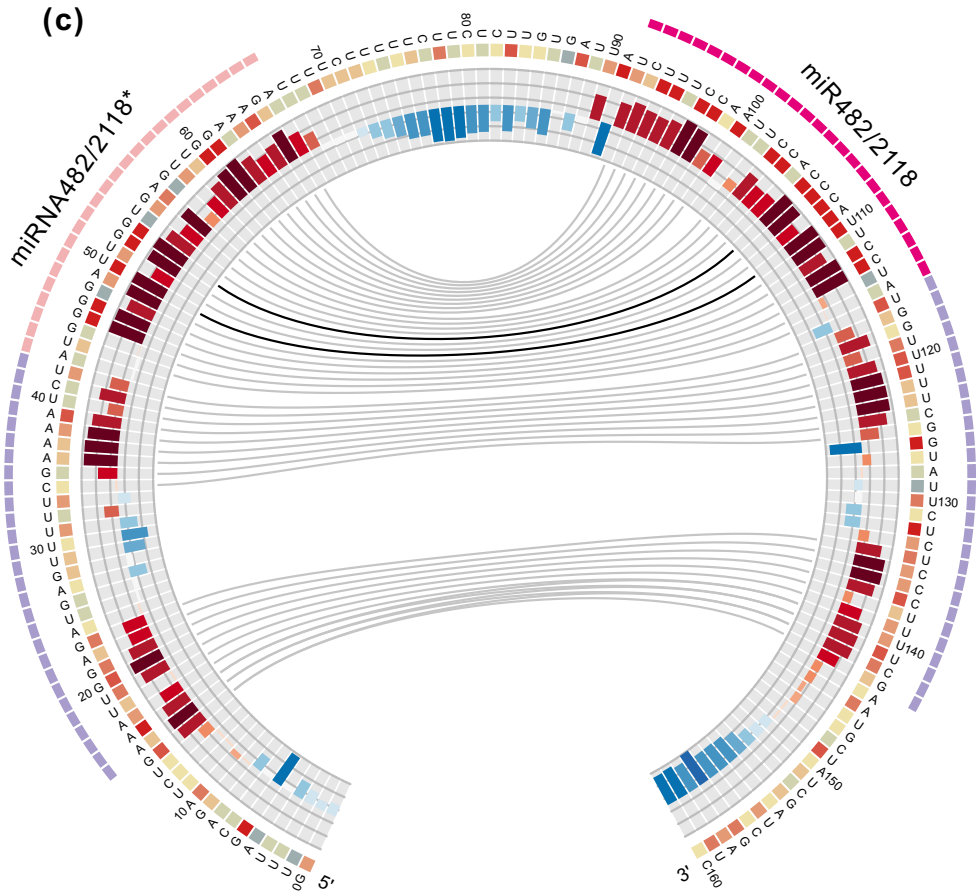
**Fig. 1 Biogenesis and conservation of miR481/2118 superfamily in plants. (a) Biogenesis and function pathway of miR482/2118.** The primary miRNA (pri-miRNA) is transcribed from the *MIRNA* gene by DNA Polymerase II. The pri-miRNA is processed by DCL1 into a precursor miRNA (pre-miRNA) and subsequently forms a duplex of miR482/2118, which is then methylated by HEN1 at the 3' terminal. With the aid of HASTY, the duplex is then translocated from nucleus to cytoplasm, in which the passenger strand is degraded while the mature strand is loaded into AGO1, consisting of RISC, to silence target genes. Typically, after the miR482/2118-directed cleavage of its target genes, the 3' cleaved fragment is converted by RDR6 into double-stranded RNAs which are processed into phasiRNAs through continuous DCL4 chopping. **(b) Structure and sequence conservation of the miR482/2118 duplex.** The nucleotide diversity of miR482/2118 and miR482/2118* are shown by sequence logos. Nucleotide pairings at each position within miR482/2118 duplex is indicated using different colors, green for A:U/C:G matches; purple for G:U matches; pink for mismatches. **(c) CIRCOS representation of sequence conservation of *MIR482/2118* precursor genes.** Sequences of ±150 bp before/after the miR/miR* region are used for the alignment. The inner ring shows a histogram of the frequency and distribution of paired (red bars), and unpaired bases (blue bars) for each position along the precursor. The outermost dotted lines denote the different sequence regions including mature (pink) and star sequences (orange), 5' and 3' lower stem regions (light blue). Gray curves refer to bases interacting with each other in the secondary structure of the precursors in >50% of the selected sequences.

**Fig. 2 Evolution and function diversification of miR482/2118. (a) Functional diversification of miR482/2118 superfamily in plant lineages.** miR482/2118 targets not only coding genes such as *NBS-LRRs* but also non-coding genes to trigger phasiRNA production in gymnosperms, and this dual function of miR482/2118 superfamily was selectivity and differentially retained in core angiosperms. (**b**) **Evolution and quantity variation of miR482/2118 and *NBS-LRR* genes in different plant species.** The numbers in the heat map represent the number of corresponding genes in different species. (**c**) **A model for the evolutionary emergence of miR482/2118 via the inverted duplication of its target *NBS-LRR* gene.** A spontaneous inverted duplication of an *NBS-LRR* gene or fragment forms a long perfect stem-loop, which can be recognized and processed into siRNA populations by DCL3 or DCL4. Over time, gradually accumulated sequence mutation results in a short and imperfect pre-miRNA-like hairpins, which is compatible for recognition and processing by DCL1, eventually turning to a canonical miRNA locus. **(d) The sequence diversity of miR482/2118 in gymnosperms, monocots and eudicots**. The sequence logos present the alignment of miR482/2118 sequences from different plant lineages. Degenerate sequence of miR482/2118 target site (the P-loop region) in *NBS-LRRs* are shown on the top, with the encoded amino acid sequences shown underneath. The nucleotide at the wobble position (the 3$^{rd}$ nucleotide of a codon) are highlighted in red and underlined. 3-nt sequences corresponding to the amino acid codon are shaded by different colors.

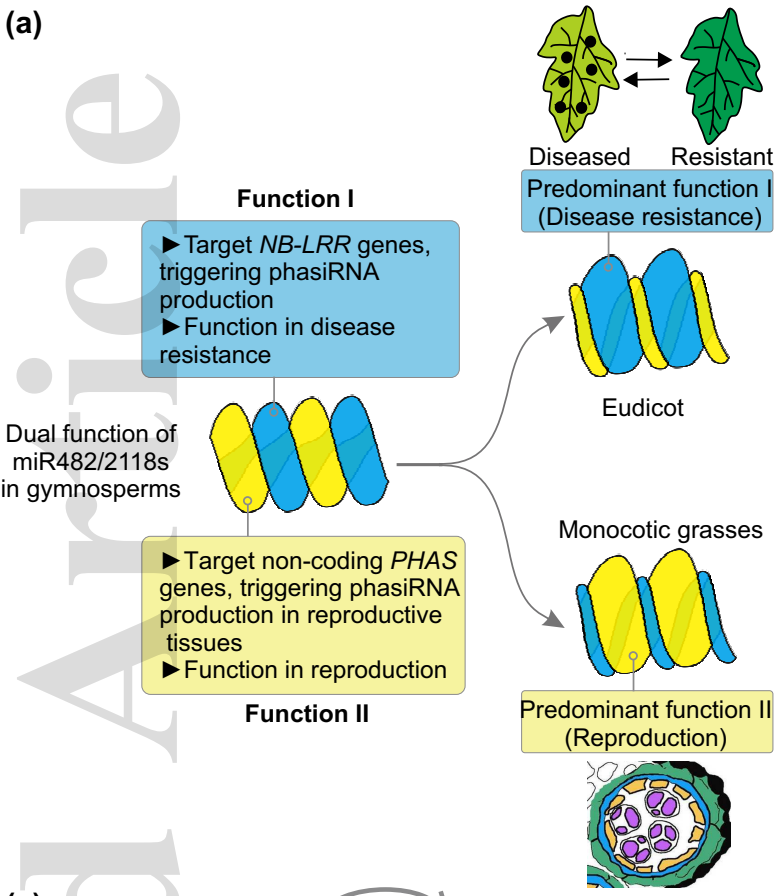
**Fig. 3 Representative functions of miR482/2118. (a) miR482/2118 targets at the conserved P-loop region of NBS-LRRs**. The NBS domain of NBS-LRR proteins contains various conserved motifs including the P-loop, Kinase-2, RNBS-B/D, and GLPL. The miR482/2118 target site is located within the P-loop region. **(b) miR482/2118 and phasiRNAs are important components in the zig-zag model for the plant innate immune response**. In this extended model, pathogen-associated molecular patterns (PAMPs) activate the PAMP-triggered immunity (PTI). The pathogen effectors may interfere with PTI resulting in effector-triggered susceptibility (ETS). Pathogen effectors may affect the production of miR482/2118, a negative regulator of *NBS-LRR* genes, to modulate the effector-triggered immunity (ETI). miR482/2118-directed phasiRNAs can reinforce the silencing effect of *NBS-LRR* genes. **(c) miR482/2118 targets long non-coding RNA (lncRNA) function in reproductive development in grasses.** Left panel represents the mechanism of *PMS1T/PMS3* regulating photoperiodic male sterility in rice. Under-long days, *PMS1T/PMS3* is targeted by miR2118, instigating subsequent phasiRNA production, which subsequently affects male fertility in rice. These phasiRNAs can *trans*-target thousands of downstream target genes. The right panel represents differential mechanisms of phasiRNAs with different 5' terminals in rice germ cells and anther wall. C-rich phasiRNAs interact specifically with MEL1 in germ cells, while U-rich phasiRNAs function via an AGO1-dependent manner in anther wall.

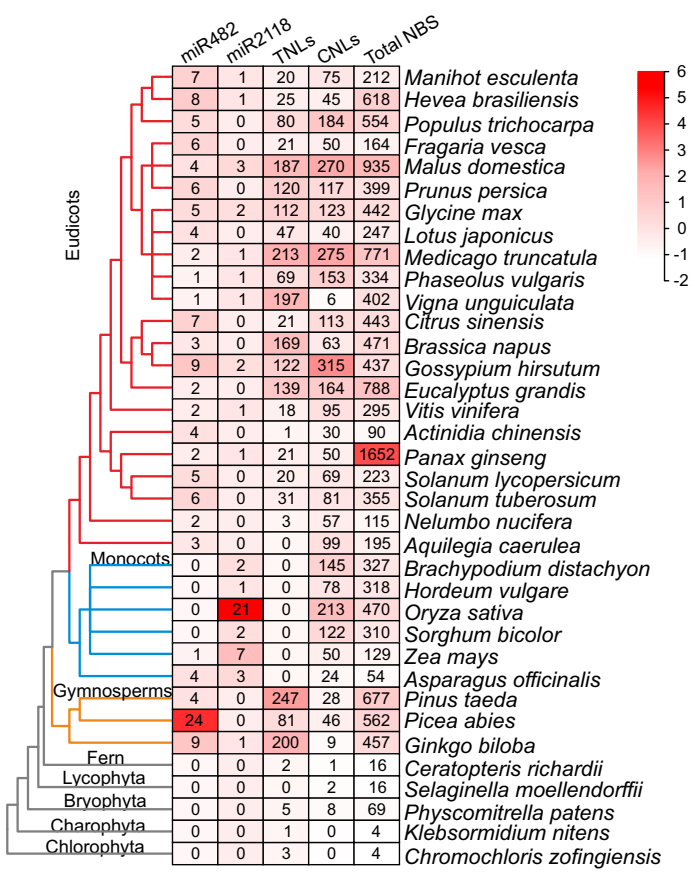


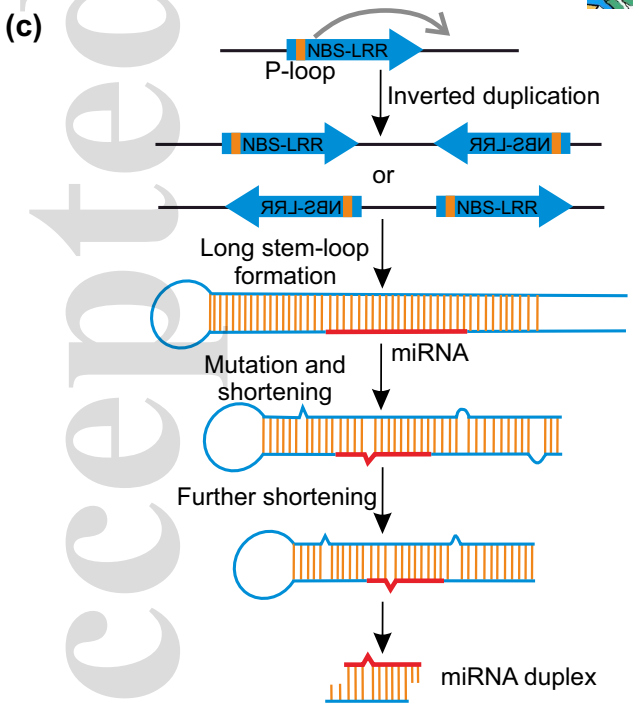



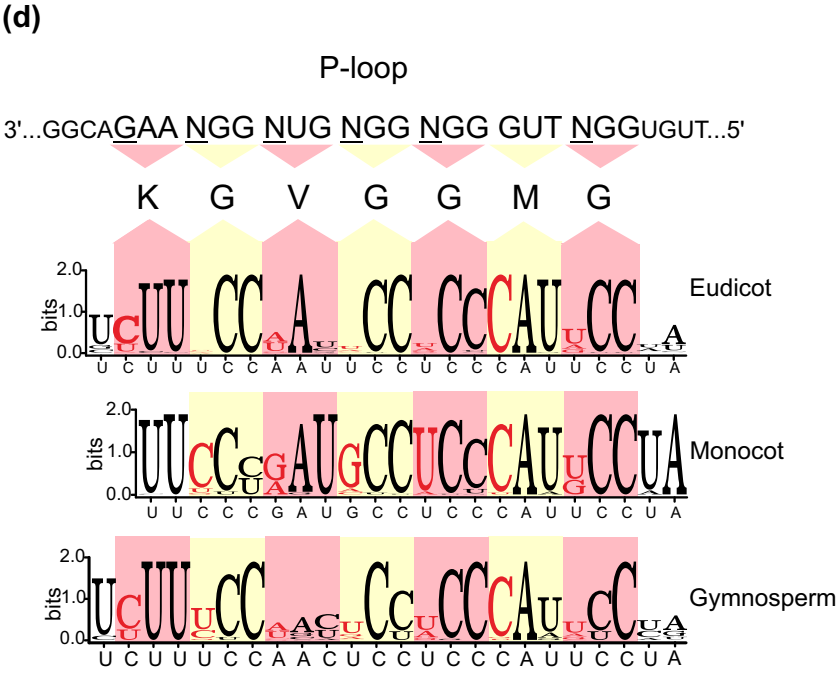




nph_17853_f1.eps

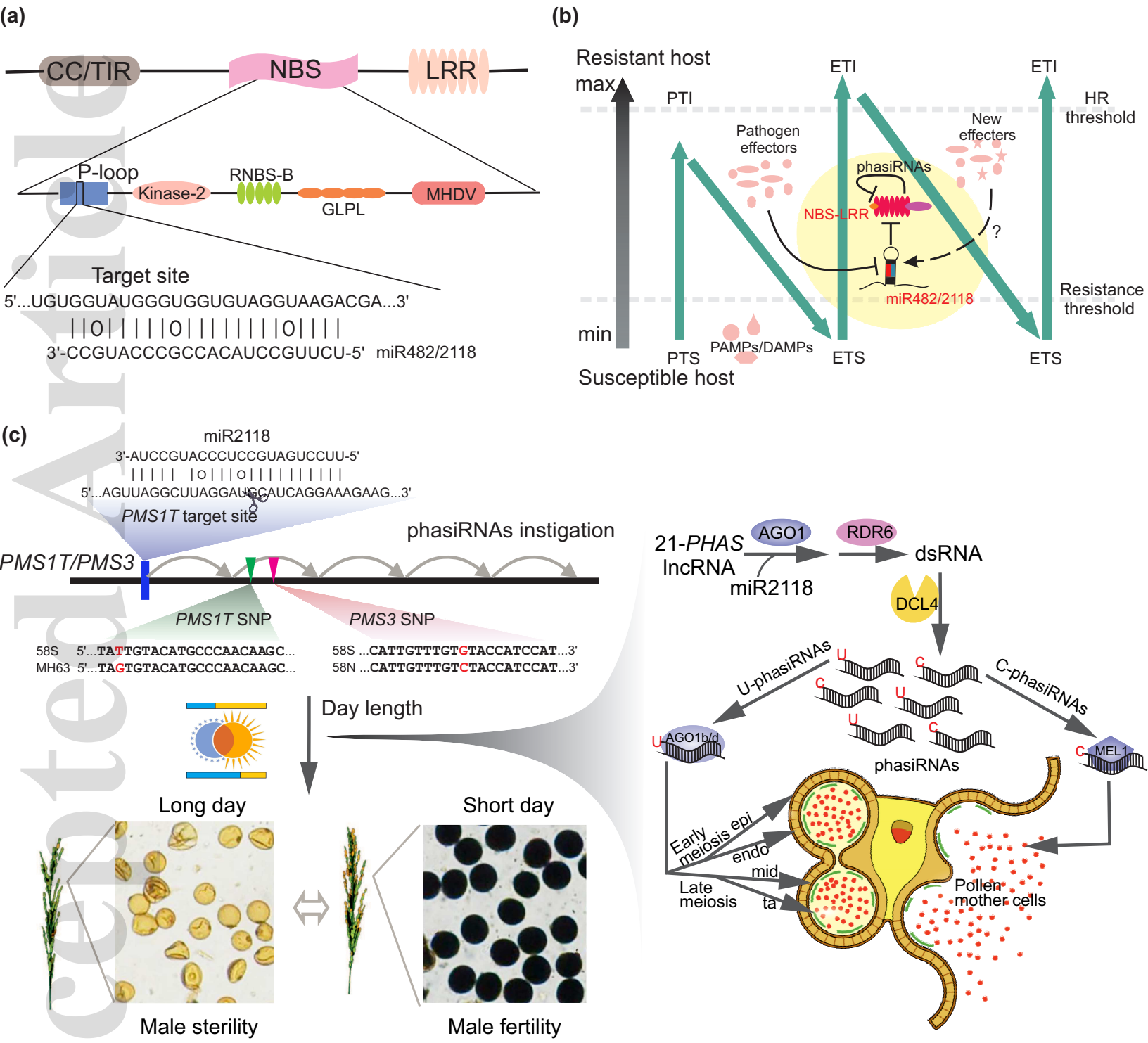



nph_17853_f2.eps




nph_17853_f3.eps